

TITLE OF THE INVENTION
IMAGE PROCESSING APPARATUS AND METHOD, AND STORAGE
MEDIUM

5 FIELD OF THE INVENTION

The present invention relates to an image processing apparatus and method and a storage medium and, more particularly, to an image processing method and apparatus for invisibly embedding information in
10 digital image data or extracting embedded information, and a storage medium.

BACKGROUND OF THE INVENTION

Conventionally, various schemes of digital
15 watermark technology have been developed as methods of protecting the copyrights of digital contents. These methods have recently received a great deal of attention as a technology for security and copyright protection in electronic distribution, in which pieces
20 of handling information of digital contents, including the copyright holder name and the buyer ID, are invisibly embedded in the digital image information, thereby enabling to track use without permission by illicit copy. A digital watermark technology as a
25 means for suppressing alteration of digital contents has also been developed. In this digital watermark

09615577-071300

technology, various data embedding methods have been proposed. In a method, information is embedded using a mask pattern. In this method, information is repeatedly embedded in digital image data in accordance with a mask pattern. For example, information is embedded at positions a, b, c, and d of each of mask patterns shown in Figs. 1A to 1D using quantization error in accordance with mask pattern arrays shown as in Figs. 2 to 4, thereby obtaining a synthetic image.

However, to improve the accuracy for specifying an altered portion in the resultant synthetic image, the mask patterns must be arrayed densely on the image data, as shown in Fig. 2 or 4. In addition, to improve watermark information detection accuracy in a partial image extracted from the synthetic image, generally, the mask pattern array as shown in Fig. 4 is preferably used. Hence, to simultaneously improve both altered portion specifying accuracy and watermark resilience against extraction, embedding is done using the mask pattern array as shown in Fig. 4.

This embedding method can improve the altered portion specifying accuracy and watermark resilience against extraction. However, this method suffers the following problems.

1) To improve the resilience, the mask pattern size is preferably as small as possible. However,

low-frequency noise or block noise becomes noticeable to degrade the image quality. In addition, the embeddable information amount is limited by the number of data in the mask pattern.

5 For example, when the mask patterns shown in Figs. 1A to 1D are used, the information amount is limited to four bits for the positions a, b, c, and d or 16 bits at maximum.

10 2) To improve the image quality, the mask pattern size is preferably as large as possible, though the resilience becomes poor.

15 3) To improve the altered portion specifying accuracy, the number of embedding positions must be large. However, low-frequency noise or block noise becomes noticeable to degrade the image quality.

20 That is, the image quality and the altered portion specifying accuracy/watermark resilience against extraction have tradeoff relationships. If one is improved, the other degrades: both cannot be simultaneously improved.

25 A digital watermark information embedding method called a patchwork method is known. In this method, the values of one part of an image are intentionally increased while the values of the other part are intentionally decreased. Hence, certain additional information can be embedded while the values of the

entire image are kept almost unchanged.

Although it is conventionally known that information must be embedded undetectably for the human eye, the method of determining the image embedding position in the above patchwork method or the like has not been established yet.

To embed digital watermark information by partially modulating an image, for example, a method of determining the modulation position at random is available. However, with this method, the image quality cannot be kept sufficiently high.

SUMMARY OF THE INVENTION

The present invention has been in consideration of the above prior art, and has as its object to embed digital watermark information by partially changing an image such that degradation in image quality is possibly unnoticeable to the human eye.

In order to achieve the above object, an image processing apparatus according to the present invention has, e.g., the following arrangement.

More specifically, according to the present invention, there is provided an image processing apparatus for embedding predetermined information in image data, comprising:

generation means for generating a mask pattern

which has a blue noise characteristic and specifies a target embedding position in an $M \times N$ size; and

embedding means for applying the mask pattern to part of the image data and modulating image data

5 corresponding to the target embedding position to embed the predetermined information.

There is also provided an image processing apparatus comprising generation means for binarizing each coefficient of a mask and generating a

10 two-dimensional mask having periodical or pseudo-periodical peaks on a radial frequency domain of resultant binary information, first input means for inputting image data, second input means for inputting additional information, means for making each
15 coefficient of the two-dimensional mask correspond to each bit information of the additional information, and digital watermark embedding means for adding/subtracting the image data on the basis of a positional relationship obtained by assigning the
20 two-dimensional mask onto the image data as a correspondence result, thereby embedding each bit information in the image data.

Other features and advantages of the present invention will be apparent from the following
25 description taken in conjunction with the accompanying drawings, in which like reference characters designate

the same or similar parts throughout the figures thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A to 1D are views showing examples of
5 target embedding regions in a digital watermark
technology;

Fig. 2 is a view showing a mask pattern array
used for information embedding;

Fig. 3 is a view showing another mask pattern
10 array used for information embedding;

Fig. 4 is a view showing still another mask
pattern array used for information embedding;

Fig. 5 is a view showing a mask pattern in the
first embodiment;

Fig. 6 is a view showing an information embedding
15 region based on the mask pattern in the first
embodiment;

Fig. 7 is a view showing the embedding sequence
in the first embodiment;

Fig. 8 is a view showing the concept of
20 information embedding in the first embodiment;

Figs. 9A and 9B are views showing an example of
an embedded image;

Figs. 10A and 10B are views showing an example of
25 an altered image;

Fig. 11 is a view showing a mask pattern

application sequence in embedded information extraction processing;

Fig. 12 is a view showing the stored state of collected information;

5 Fig. 13 is a view for explaining processing of extracting embedded information;

Fig. 14 is a block diagram showing an apparatus according to the first embodiment;

10 Fig. 15 is a flow chart showing embedding processing in the first embodiment;

Fig. 16 is a flow chart showing embedded information extraction/alteration determination processing in the first embodiment;

15 Fig. 17 is a block diagram showing the overall arrangement of a digital watermark embedding apparatus according to the second embodiment;

Fig. 18 is a block diagram showing the overall arrangement of a digital watermark extraction apparatus;

20 Fig. 19 is a view showing an example of image data generated on the extraction side in print system processing;

Fig. 20 is a block diagram showing a registration signal embedding means;

25 Fig. 21 is a view for explaining a registration signal;

Fig. 22 is a flow chart showing the processing contents of a reliability distance calculation means;

Fig. 23 is a block diagram showing a scale matching means;

5 Figs. 24A and 24B are graphs for explaining registration signal extraction;

Fig. 25 is a view showing a mask pattern array used to embed and extract additional information;

10 Fig. 26 is a flow chart showing the processing contents of an additional information embedding means;

Fig. 27 is a block diagram showing an embedding position determination means;

15 Fig. 28 is a graph showing the appearance frequency distribution of coefficient values of a cone mask or blue noise mask;

Fig. 29 is a graph showing the radial frequency characteristic of the human eye;

20 Figs. 30A and 30B are graphs showing the radial frequency characteristics of the blue noise mask and cone mask, respectively;

Fig. 31 is a view for explaining a position reference mask;

Fig. 32 is a view showing embedding positions in the position reference mask;

25 Figs. 33A and 33B are views showing a state wherein the pattern array is bitmapped on the mask

shown in Fig. 32;

Figs. 34A and 34B are views showing a region necessary for embedding additional information Inf in the entire image;

5 Fig. 35 is a view for explaining calculations for embedding of the additional information Inf;

Fig. 36 is a block diagram for explaining an additional information extraction means;

10 Fig. 37 is a view for explaining a state wherein the additional information Inf is extracted;

Fig. 38 is a view showing a state wherein the additional information Inf is tried to be extracted although it is not present;

15 Fig. 39 is a graph showing an ideal appearance frequency distribution when a reliability distance d is extracted from the original image;

Fig. 40 is a graph showing a case wherein the reliability distance d is extracted from an image with a digital watermark embedded;

20 Fig. 41 is a graph for explaining examples of the appearance frequency distribution of reliability distances d1 and d2 in the second embodiment;

Fig. 42 is a view for explaining the principle of registration signal embedding and extraction;

25 Fig. 43 is a view showing offset matching processing;

Fig. 44 is a flow chart for explaining offset matching processing;

Fig. 45 is a block diagram showing a registration signal embedding means in a spatial region;

5 Fig. 46 is a view for explaining two sets in a patchwork method;

Fig. 47 is a flow chart for explaining the entire contents of digital watermark embedding processing;

10 Fig. 48 is a flow chart for explaining the entire contents of digital watermark extraction processing;

Figs. 49A and 49B are views showing examples of a pattern array perpendicular to the pattern shown in Fig. 25;

15 Fig. 50 is a view for explaining patterns "perpendicular" to each other;

Figs. 51A and 51B are views showing first and second position reference masks;

Fig. 52 is a view showing the structure of the additional information Inf;

20 Fig. 53 is a view showing examples of coefficients in the blue noise mask; and

Fig. 54 is a view showing an example of the coefficients of pixel values of a cone mask.

25 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The embodiments of the present invention will be

described below in detail with reference to the accompanying drawings.

<First Embodiment>

Fig. 14 is a block diagram showing an information processing apparatus according to the first embodiment. Referring to Fig. 14, a CPU 1 controls the entire apparatus. A ROM 2 stores a BIOS and boot program. A RAM 3 in which an OS and various applications are loaded is also used as a work area of the CPU. In the RAM 3, an image memory 3a for storing a read image is ensured. An external storage device 4 such as a hard disk stores the OS and various application programs. Processed image data can also be stored in the external storage device 4 as files. The apparatus also has a keyboard 5 (including a pointing device such as a mouse). An image scanner 6 for reading an original image as color image data is connected through an interface such as a SCSI. A display control section 7 includes a display memory which stores image data to be displayed. A display device 8 such as a CRT or liquid crystal display displays a video signal output from the display control section 7. A color printer 9 has a printer engine having printheads for discharging ink droplets by thermal energy in units of print color components. A network communication section 10 communicates with a network (may be the Internet).

In this embodiment having the above arrangement, color image data read by the image scanner is stored in the image memory 3a, and independently set information is embedded using the digital watermark technology.

5 Fig. 5 is a view showing a mask pattern having 32 × 32 pixels, which is comprised of intermediate- and high-frequency band components excluding low-frequency components, i.e., has a so-called blue noise characteristic. A black dot in Fig. 5 indicates an
10 information embedding position. The amount (the number of bits) of information to be embedded is equal to or smaller than the number of black dots.

The mask pattern having the blue noise characteristic is normally used to binarize an image.
15 In this embodiment, information is embedded using the pattern having the blue noise characteristic, thereby efficiently performing digital watermark processing in consideration of human visual characteristics. The structure of a blue noise mask pattern is disclosed in,
20 e.g., Robert Ulichney, "Digital Halftoning", Massachusetts Institute of Technology (1987) or Japanese Patent No. 2622429.

Fig. 6 is a view showing the layout of mask patterns, which is used to embed information in entire
25 image data using the mask pattern shown in Fig. 5. A method of embedding and detecting information using the

Conversely, to detect embedded data from an image, basically, if the pixel value is an even multiple (0, 16, 32,...) of quantization step width, data that has been embedded at that pixel position can be detected as "0". If the pixel value is an odd multiple (8, 24,...), the embedded data can be detected as "1". However, this applies to only a pixel which is determined to be at the target embedding position. For data with another value (e.g., 10), information is at least not embedded at that pixel position.

Assume that there are image data (Fig. 9A) converted by embedding, partially extracted image data (Fig. 9B), image data (Fig. 10A) obtained by altering the image data shown in Fig. 9A, and image data (Fig. 10B) obtained by partially extracting the image data shown in Fig. 9A and altering the image data.

In this case, as shown in Fig. 11, starting from the upper left corner of each image, dots represented in black in the mask pattern shown in Fig. 5 are checked to determine whether information is embedded. At this time, the coordinates of a point (x,y) in Fig. 11 and the number of pixels which are determined to have information embedded are recorded.

Whether information is embedded in a dot (pixel) is determined by determining whether the value of each pixel masked by the black dot of the mask pattern shown

in Fig. 5 is an integer multiple of quantization step width. The number of pixels (to be referred to as a determination count hereinafter) whose values are determined to be integer multiples of quantization step width, and the point (x,y) representing the mask pattern position at that time are stored in an appropriate area in the RAM 3 in correspondence with each other.

This processing is repeated while moving the mask pattern to the right by one pixel. When the mask pattern has reached the right end, it is moved to a position shifted from the upper left corner in Fig. 11 to the lower side by one dot, and the same processing as described above is repeated.

When determination processing for the entire image to be checked is ended, the values are sorted in descending order of determination counts. The coordinates and determination count of the point (x,y) with a determination count equal to or larger than a threshold value are obtained. As the threshold value, the embedded information amount (the number of bits) is used. When q bits are required for an author name, and the amount of information to be embedded is Q bits (the number of black dots of the mask pattern $> Q > q$), the remaining bits (Q - q) contain an appropriate value such as a parity.

09615577.071300

Fig. 12 is a view showing k sets of information obtained in the above way (the number of information equal to or larger than the threshold value is k).

For each of the sets of information, the embedded
5 information is determined in units of mask pattern embedding positions. At this time, determination information corresponding to the embedding position of each set is recorded.

Fig. 13 is a view showing pieces of determination
10 information corresponding to the embedding positions P_1 to P_n of the respective sets. According to be above-described determination rule, when the value of pixel data X is an even multiple of quantization step width h , it is determined that information "0" is
15 embedded, and when the pixel value is an odd multiple of quantization step width h , it is determined that information "1" is embedded.

The embedded information is determined by
decision by majority for "0" and "1" at the same
20 embedding position.

An altered portion is specified in the following way.

The x - and y -coordinate values of the point in the majority depend on the mask pattern size (32×32
25 in this embodiment). More specifically, without alteration, when $x = 0$ and $y = 0$ are set for the upper

left corner position of the input image, x and y of the point (x,y) are basically given by

$$x = 32 \times i + c1$$

$$y = 32 \times j + c2$$

- 5 for i and $j = 0, 1, 2, \dots$, and $c1$ and $c2$ are constants (depending on the input image). Since the information is embedded in the entire image, i and j are originally consecutive.

- 10 Assume that x and y of the point (x,y) have the above relationships, and when $j = 5$, a region represented by $i = 1, 2, 3, 4, 10, 11$ is determined as a target embedding region, it can be determined that regions represented by $j = 5$ and $i = 5$ to 9 are altered.

- 15 When the altered portion is determined, a message representing the alteration and the altered portion in the input image data are displayed on the display device 8 such that the altered portion can be discriminated from, e.g., an unaltered portion. As an example of discriminative display, the altered portion
20 is enclosed with a frame or displayed in another color.

In this embodiment, information is embedded using a square mask pattern as shown in Fig. 6. However, a mask pattern can generally have an $M \times N$ size.

- 25 In this embodiment, the amount (the number of bits) of information to be embedded must not exceed the number of black dots of the mask pattern. To store an

In step S26, it is determined whether processing has reached the right end of the image. If NO in step S26, the value x is incremented by "1", i.e., the mask pattern application position is shifted by one pixel,
5 and the flow returns to step S24.

If YES in step S26, the value x is initialized to "0", and the value y is incremented by "1" in step S28. Processing in steps S24 to S28 is repeated until it is determined in step S29 that processing in steps S24 and
10 S25 is complete for the entire image.

When the pieces of information corresponding to one frame are collected, the stored data are arranged in descending order of counts, and data whose counts are equal to or larger than a predetermined value are
15 validated in step S30. The flow advances to step S31 to extract the embedded information.

In step S32, on the basis of the pieces of collected information and the values x and y of each information, it is determined whether the image is
20 altered, and if so, the altered portion is determined. If it is determined that the image is not altered, the embedded information is displayed, e.g., the copyright holder name is displayed in step S33.

If it is determined that the image is altered,
25 error processing is performed to display a message representing the alteration and explicitly indicate the

altered portion in step S34, and the processing is ended.

As described above, according to the first embodiment, a mask pattern having no low-frequency components is used to embed information. Also, instead of embedding information at black dots of the mask pattern, i.e., all positions at "1", quantization and information embedding are performed at limited positions. Hence, the influence on the image quality can be reduced, and a satisfactory image quality can be maintained. Additionally, even when the image is altered, the altered portion can be specified.

The present invention may be applied to a single
apparatus or a system constituted by a plurality of
15 apparatuses.

Although the above embodiment requires, e.g., a means for inputting an image (a means for connecting an image scanner or network, or hardware such as a floppy disk), this means can be a general-purpose device normally incorporated in or connectable to a general-purpose information processing apparatus (personal computer), and its processing can be realized by CPU processing, i.e., a program.

Hence, the present invention can be implemented
25 even by supplying a storage medium storing software
program codes for realizing the functions of the

above-described embodiment to a system or apparatus, and causing the computer (or a CPU or MPU) of the system or apparatus to read out and execute the program codes stored in the storage medium.

5 In this case, the program codes read out from the storage medium realize the functions of the above-described embodiments by themselves, and the storage medium storing the program codes constitutes the present invention.

10 As a storage medium for supplying the program codes, a floppy disk, a hard disk, an optical disk, a magnetooptical disk, a CD-ROM, a CD-R, a magnetic tape, a nonvolatile memory card, a ROM, or the like can be used.

15 The functions of the above-described embodiment are realized not only when the readout program codes are executed by the computer but also when the OS (Operating System) running on the computer performs part or all of actual processing on the basis of the instructions of
20 the program codes.

 The functions of the above-described embodiment are also realized when the program codes read out from the storage medium are written in the memory of a function expansion board inserted into the computer or a
25 function expansion unit connected to the computer, and the CPU of the function expansion board or function

expansion unit performs part or all of actual processing on the basis of the instructions of the program codes.

As described above, according to the present invention, in the digital watermark technology,

5 degradation in image quality can be reduced, and the altered portion specifying accuracy can be improved.

<Second Embodiment>

[1 Digital Watermark Embedding Apparatus]

The outline of a digital watermark embedding
10 apparatus according to the second embodiment will be described below with reference to the accompanying drawings.

Fig. 17 shows the digital watermark embedding apparatus of the second embodiment. As shown in
15 Fig. 17, the digital watermark embedding apparatus comprises a color component extraction means 101, registration signal embedding means 102, embedding position determination means 103, additional information embedding means 104, and color component
20 synthesis means 105.

Image data I is input to the digital watermark embedding apparatus. This image data is a multilevel image data in which predetermined bits are assigned to each pixel. In this embodiment, the input image data I
25 may be either grayscale image data or color image data. Grayscale image data has one element per pixel. Color

image data has three elements per pixel. In this embodiment, the three elements are red, blue, and green components. However, the present invention can also be applied to a combination of different color components.

5 The image data I input to the digital watermark embedding apparatus is input to the color component extraction means 101 first.

 When the input image data I is color image data, the color component extraction means 101 separates only
10 the blue component from the color image data and outputs the component to the registration signal embedding means 102 on the output side.

 The remaining color components are output to the color component synthesis means 105 on the output side.
15 That is, only a color component in which digital watermark information is to be embedded is separated and sent to the digital watermark processing system.

 In this embodiment, digital watermark information is embedded in the blue component. This is because the
20 human eye is most insensitive to the blue component in the red, blue, and green components. Hence, when digital watermark information is embedded in the blue component, degradation in image quality due to the digital watermark information can hardly be perceived
25 by the human eye, unlike a case wherein the digital watermark information is embedded in another color

The image data of the blue component obtained by the color component extraction means 101 is input to the registration signal embedding means 102. The registration signal embedding means 102 embeds a registration signal in the image data using a kind of digital watermark technology. That is, the human eye cannot perceive the registration signal embedded in the image data. The method of embedding the registration signal will be described later in detail.

10 The registration signal embedding means 102 outputs the image data with the registration signal embedded.

The embedding position determination means 103 determines the embedding position of additional information Inf in the image data input from the registration signal embedding means 102.

The embedding position determination means 103 outputs control data representing the embedding position of the additional information Inf in the image to the additional information embedding means 104 together with the input image data.

The additional information embedding means 104 receives the additional information Inf (a plurality of bit information) in addition to the image data and control data. The additional information Inf is embedded at the determined embedding position in the

image data of the blue component using the digital watermark technology. Embedding of the additional information Inf using the digital watermark technology will also be described later.

5 The additional information embedding means 104 outputs the image data with the additional information Inf embedded. The image data is input to the color component synthesis means 105.

10 The color component synthesis means 105 synthesizes the blue component processed on the input side (to the additional information embedding means 104) and the red and green components directly input from the color component extraction means 101 into normal color image data.

15 With the above processing, image data wI in which the registration signal and additional information Inf are embedded by the digital watermark technology is output.

20 In this embodiment, a description will be made assuming that attacks for generating various geometrical distortions are made against the image data wI. For example, the image is intentionally edited by the user, or after the image data wI is printed, the print is scanned with a scanner. Image data wI' shown
25 in Fig. 18 is the attacked image data.

 The overall flow by the above-described means

processing will be described later in detail with
reference to Fig. 26. A macro block means a minimum
embedding unit. One complete additional information
Inf is completely embedded in an image region
5 corresponding to a macro block.

After the additional information Inf is embedded
in the image data, the image data wI with the digital
watermark information embedded is output in step 3409.

As described above, attacks for generating
10 various geometrical distortions may be made against the
image data wI before digital watermark extraction start
processing in Fig. 48 (to be described later) is
executed.

[2 Digital Watermark Extraction Apparatus]

15 The outline of a digital watermark extraction
apparatus according to the second embodiment will be
described next.

Fig. 18 is a block diagram showing the digital
watermark extraction apparatus according to the second
20 embodiment. As shown in Fig. 18, the digital watermark
extraction apparatus comprises a color component
extraction means 201, registration means 202, and
additional information extraction means 203.

The image data wI' is input to the digital
25 watermark extraction apparatus. The image data wI' may
have received attacks against the image data wI to

generate various geometrical distortions. The attacks include irreversible compression such as JPEG compression, scaling, rotation, printing & scanning, and a combination thereof.

5 Although the image data wI' and wI ideally have the same contents, actually, the two image data often have considerably different contents.

09615577.071300
10 The color component extraction means 201 receives the image data wI' , extracts the blue component, and outputs the image data of the blue component to the registration means 202 on the output side. The red and green components of the image data wI' , except the blue component, are unnecessary and therefore are discarded.

15 The registration means 202 receives image data wI_2' of the blue component obtained by the color component extraction means 201. Using the image data wI_1' of the blue component, image data wI_2' whose geometrical distortions are corrected is generated.

20 As described above, the image data wI' may have a scale different from that of the image data wI . However, the image data wI_2' always has the same scale as that of the image data wI . The reason for this and processing of equalizing the scale of the image data wI_2' to that of the image data wI will be described
25 later in detail.

 The registration means 202 outputs the image data

in step 3505. The embedded additional information Inf is extracted from the image data wI_2' whose scale and offset are already corrected.

In statistical authorization step 3507, the accuracy of the extracted digital watermark information is calculated and determined. If it is determined that the digital watermark information is incorrect, the flow returns to step 3502 to re-input an image which is supposed to have digital watermark information. If it is determined that the digital watermark information is sufficiently correct, the digital watermark information (additional information Inf) is extracted by comparison processing in step 3508. In step 3510, the information representing the accuracy is displayed as a reliability index D (to be described later).

[3 Detailed Description of Each Section]

Each section will be described next in detail.

Registration processing executed in step 3503 by
the registration means 202 on the digital watermark
20 extraction side will be described first.

Registration processing is pre-processing of digital watermark information extraction, which is executed to enable digital watermark information extraction from the image data wI' input to the digital watermark extraction apparatus. First, changes that may occur in image data processed by a printing system

will be considered below. Registration processing for such changes will be examined, and registration processing for the printing system will be considered.

The digital watermark extraction apparatus does
5 not always directly receive the image data wI output
from the digital watermark embedding apparatus.

A case wherein the image data wI is printed by a YMCK inkjet printer, and the resultant print is scanned with a scanner will be exemplified.

10 If the output resolution of the printer is
different from the input resolution of the scanner, the
image data obtained by scanning has a scale different
from that of the original color image data wI. Digital
watermark information can be accurately extracted from
15 the obtained image data wI' with a low possibility.
Hence, a means capable of correcting the difference in
scale must be prepared.

In this embodiment, since both the input resolution and output resolution are known, the scale ratio can be calculated. For example, when the output resolution is 600 dpi, and the input resolution is 300 dpi, the scale ratio of the image before printing to that after scanning is 2. In accordance with the calculated scale ratio, scaling is performed for the image data wI' using an appropriate scaling algorithm. With this processing, the image sizes of the image data

by intentional editing.

The registration signal embedding means and registration means which are provided to solve the above problem when the ratio of input and output
5 resolutions is unknown will be described below.

[3-1 Registration Signal Embedding Processing]

The registration signal embedding means 102 (step 3404) will be described first in detail.

The registration signal embedding means 102 is
10 located on the input side of the additional information embedding means 104. This means 102 embeds, in the original image data in advance, a registration signal to be referred to for registration of the image data
WI' by the registration means shown in Fig. 18. The
15 registration signal is hard to perceive with the human eye as digital watermark information and embedded in the image data (blue component of color image data in this embodiment).

Fig. 20 is a block diagram showing the internal
20 arrangement of the registration signal embedding means 102. The registration signal embedding means 102 comprises a block segmentation means 401, Fourier transform means 402, addition means 403, inverse Fourier transform means 404, and block synthesis means
25 405 shown in Fig. 20. Each means will be described below in detail.

09645577.071300

The block segmentation means 401 segments the input image data into a plurality of blocks which do not overlap each other. In this embodiment, the block size is set to a power of 2. Actually, another size
5 may be used. When the block size is a power of 2, the Fourier transform means 402 following the block segmentation means 401 can perform high-speed processing.

The blocks segmented by the block segmentation
10 means 401 are divided into two sets I_1 and I_2 . The set I_1 is input to the Fourier transform means 402 on the output side while the set I_2 is input to the block synthesis means 405 on the output side. In this embodiment, as the set I_1 , one of the blocks obtained
15 by the block segmentation means 401, which is located closest to the center of the image data I , is selected. All the remaining blocks are selected as the set I_2 .

This is because this embodiment can be implemented using at least one block, and a smaller
20 number of blocks shorten the processing time. However, the present invention is not limited to this and also incorporates a case wherein two or more blocks are selected as the set I_1 .

The digital watermark embedding apparatus and
25 digital watermark extraction apparatus must share the information of block size and blocks to be selected as

a registration signal embedding target.

The part I_i of the image data obtained by segmentation by the block segmentation means 401 is input to the Fourier transform means 402.

5 The Fourier transform means 402 executes Fourier transform for the input image data I_i . The original data form of the input image data I_i is called a spatial domain while the data form after Fourier transform is called a frequency domain. Fourier
10 transform is executed for all the input blocks. In this embodiment, since the size of the input block is a power of 2, fast Fourier transform is used to increase the processing speed.

 Although Fourier transform requires a calculation
15 amount for $n \times n$ times, fast Fourier transform is a transform algorithm which can be executed in a calculation amount $(n/2)\log_2(n)$ (n is a positive integer). Fast Fourier transform and Fourier transform
20 are different only in the speed for obtaining the calculation result, and the same result is obtained by these calculations. Hence, in the description of this
 embodiment, fast Fourier transform and Fourier transform are not discriminated.

 Image data in the frequency domain obtained by
25 Fourier transform is represented by an amplitude spectrum and phase spectrum. Only the amplitude

spectrum is input to the addition means 403. On the other hand, the phase spectrum is input to the inverse Fourier transform means 404.

The addition means 403 will be described next.

5 The addition means 403 receives a signal r called a registration signal as well as the amplitude spectrum. An example of the registration signal is an impulse signal as shown in Fig. 21.

Fig. 21 is a view showing the amplitude spectrum
in two-dimensional radial frequency components obtained
by Fourier transform. A low-frequency component is at
the center, and high-frequency components are at the
periphery. An amplitude spectrum 501 is the amplitude
spectrum of signal components of the original image
components. In a signal corresponding to a natural
image such as a photograph, many large signals
concentrate to the low-frequency region. On the other
hand, almost no signals are present in the
high-frequency region.

20 In this embodiment, a description will be made
assuming that a series of processing operations are
executed for a natural image. However, the present
invention is not limited to this, and a document image
or CG image can also be processed in the same way.
25 However, this embodiment is especially effective in
processing a natural image having a relatively large

number of halftone components.

Fig. 21 shows an example of this embodiment in which impulse signals 502, 503, 504, and 505 are added to the horizontal/vertical Nyquist frequency components in the frequency domain of the original signal 501 of the natural image. As shown in this example, the registration signal is preferably an impulse signal. This is because only the registration signal can be easily extracted by the digital watermark extraction apparatus to be described later.

Although impulse signals are added to the Nyquist frequency components of the input signal in Fig. 21, the present invention is not limited to this. More specifically, any other signal can be used as far as the registration signal is not removed when the image with digital watermark information embedded has received an attack. As described above, irreversible compression such as JPEG compression has the low-pass filter effect. Hence, even when an impulse signal is embedded in a high-frequency component as an information compression target, the signal may be removed by compression/expansion processing.

On the other hand, when an impulse is embedded in a low-frequency component, the signal is readily perceived as noise due to the human visual characteristics, as compared to embedding in a

high-frequency component. Hence, in this embodiment, the impulse signal is embedded in a frequency of intermediate level higher than the first frequency with which the signal is hardly perceived by the human eye and lower than the second frequency with which the signal is hardly removed by irreversible compression/expansion processing. This registration signal is embedded in each of blocks (one block in this embodiment) input to the addition means 403.

10 The addition means 403 outputs the signal obtained by adding the registration signal to the amplitude spectrum of the image data in the frequency domain to the inverse Fourier transform means 404.

15 The inverse Fourier transform means 404 executes inverse Fourier transform for the input image data in the frequency domain. This inverse Fourier transform is executed for all the input blocks. As in the Fourier transform means 402, since the size of the input block is a power of 2, fast Fourier transform is used to increase the processing speed. The signal in the frequency domain input to the inverse Fourier transform means 404 is converted into a signal in the spatial domain by inverse Fourier transform and output.

25 The image data in the spatial domain output from the inverse Fourier transform means 404 is input to the block combining means 405.

signal r' by inverse Fourier transform. The registration signal r is a signal on the frequency domain, like that shown in Fig. 21. The block from the block segmentation means 3201 and the signal r' from the inverse Fourier transform means 3204 are input to the block synthesis means 3203 and added. The signal output from the addition means 3202 is input to the block synthesis means 3203. The image data (blue component) is reconstructed and output.

10 The means shown in Fig. 45 performs the same processing as that by the means shown in Fig. 20 in the spatial domain. Since no Fourier transform means is required, unlike the means shown in Fig. 20, high-speed processing is possible.

15 Referring to Fig. 45, the signal r' is a signal independent from the input image data I . Hence, the signal r' can be generated in advance instead of calculating the signal r' , i.e., executing processing by the inverse Fourier transform means 3204 every time input image data I is input. In this case, the registration signal can be embedded at a higher speed by omitting the inverse Fourier transform means from the means shown in Fig. 45. Registration processing of referring to the registration signal will be described later.

<<Patchwork Method>>

This embodiment uses a principle called a patchwork method to embed the additional information Inf. The principle of patchwork method will be described first.

5 In the patchwork method, the additional information Inf is embedded by generating a statistical bias.

This will be described with reference to Fig. 46. Referring to Fig. 46, reference numerals 3301 and 3302
10 denote subsets of pixels; and 3303, an entire image. Two subsets A 3301 and B 3302 are selected from the entire image 3303.

The additional information Inf can be embedded by the patchwork method of this embodiment as long as the
15 two selected subsets do not overlap each other. However, the size or selection method for the two subsets largely influences the resilience of additional information Inf embedded by the patchwork method, i.e., the strength for preventing the additional information
20 Inf from missing when an attack is made against the image data wI. This will be described later.

Let $\{a_1, a_2, \dots, a_N\}$ be the value of an element of the selected subset A and $\{b_1, b_2, \dots, b_N\}$ be the value of an element of the subset B. More specifically,
25 the values $\{a_1, a_2, \dots, a_N\}$ and $\{b_1, b_2, \dots, b_N\}$ are the values of pixels (corresponding to the value of the

blue component in color image data, in this embodiment)
included in the subsets.

An index d is defined.

$$d = 1/N \sum (a_i - b_i)$$

- 5 This value represents the expectation value of
the difference in pixel value between the two sets.

When, for a general natural image, an appropriate
subset A and subset B are selected, and the index d is
defined,

10 $d \cong 0$

The index d will be referred to as a reliability
distance d hereinafter.

- On the other hand, as an operation of embedding
each bit of the additional information Inf, operations
15 represented by

$$a'_i = a_i + c$$

$$b'_i = b_i - c$$

- are performed. These are operations of adding a value
 c to all elements of the subset A and subtracting the
20 value c from all elements of the subset B.

As in the above-described case, the subset A and
subset B are selected from the image with the
additional information Inf embedded, and the index d is
calculated.

- 25 Then,

$$d = 1/N \sum (a'_i - b'_i)$$

$$\begin{aligned}
&= 1/N \sum \{ (a_i + c) - (b_i - c) \} \\
&= 1/N \sum (a_i - b_i) + 2c \\
&\doteq 2c
\end{aligned}$$

The index d is not 0.

5 More specifically, the reliability distance d is calculated for a given image. If $d \cong 0$, it can be determined that the additional information Inf is not embedded. If the reliability distance d has a value separated from 0 by a predetermined amount or more, it
10 can be determined that the additional information Inf is embedded.

The basic concept of patchwork method has been described above.

Using this principle of patchwork method, a
15 plurality of bit information are embedded in this embodiment. In this method, the method of selecting the subset A and subset B is also defined by a pattern array.

In the above-described method, the additional
20 information Inf is embedded by adding or subtracting the element of a pattern array to or from a predetermined element of the original image.

Fig. 25 is a view showing a simple example of the pattern array. Fig. 25 shows a pattern array used to
25 refer to 8 × 8 pixels in embedding one bit, which indicates the amounts of changes in pixel values from

is laid out in different regions of the original image data a plurality of number of times to increase/decrease the pixel values, thereby embedding a plurality of bit information, i.e., additional
5 information Inf. In other words, not only the combination of the subsets A and B but also a plurality of combinations including the combination of subsets A' and B', combination of A" and B",... are assumed in different regions of one image, thereby embedding the
10 additional information Inf formed from a plurality of bits.

In this embodiment, when the original image data is large, the additional information Inf is repeatedly embedded. This is because the patchwork method uses a
15 statistical nature, and a sufficient number of information are necessary for the statistical nature to appear.

In this embodiment, to prevent regions where the pixel values are to be changed to embed a plurality of
20 bits from overlapping each other, the relative positions between the bits for use of the pattern array are determined in advance. More specifically, the relationship between the pattern array position at which the first bit information of the additional
25 information Inf is to be embedded and the pattern array at which the second bit information is to be embedded

is appropriately defined.

For example, if the additional information Inf is constructed by 16 bits, the positional relationship between the 8 × 8 pixel pattern arrays of the first to 16th bits is relatively given such that degradation in image quality is reduced in a region larger than 32 × 32 pixels.

When the image data is large, the additional information Inf (each bit information of the additional information Inf) is repeatedly embedded as many times as possible. This aims at accurately extracting each bit of the additional information Inf. Especially, in this embodiment, this repetition is important because statistical measurement is performed using the fact that the same additional information Inf is repeatedly embedded.

The above-described embedding position selection is executed by the embedding position determination means 103 shown in Fig. 17. The operation of this embedding position determination means will be described next.

[3-2 Embedding Position Determination Processing]

Fig. 27 is a block diagram showing the internal arrangement of the embedding position determination means 103.

A mask generation means 1101 shown in Fig. 27

generates a mask for defining the embedding position of each bit information of the additional information Inf. The mask is a matrix having position information for defining the relative layout of the pattern array (Fig. 25) corresponding to each bit information.

Fig. 33A shows an example of a mask 1701. Coefficient values are assigned in the mask. The coefficient values have the same appearance frequency in the mask. When this mask is used, the additional information Inf having 16 bits at maximum can be embedded.

A mask reference means 1102 loads the mask generated by the mask generation means 1101 and makes each coefficient value in the mask correspond to information representing the ordinal number of each bit information, thereby determining the pattern array layout for embedding each bit information.

A mask pattern array corresponding means 1103 bitmaps the array elements (8×8 size) of each pattern array at the position of each coefficient value in the mask. More specifically, each coefficient value (one cell) of the mask shown in Fig. 33A is extended to 8×8 times, as shown in Fig. 33B, such that the mask can be referred to as the embedding position of each pattern array.

The additional information embedding means 104

(to be described later) refers to embedding start coordinates 1702 shown in Fig. 33B and embeds each bit information using the pattern array.

In this embodiment, the mask is generated every
5 time image data (blue component) is input to the mask generation means 1101. Hence, when image data with a large size is input, the same additional information Inf is repeatedly embedded a plurality of number of times.

10 In the above method, the mask arrangement (array of coefficient values) functions as a key for embedding the additional information Inf from the image. That is, only a key holder can extract the information.

The present invention also incorporates a case
15 wherein instead of generating a mask in real time, a mask generated in advance is stored in, e.g., the internal storage means of the mask generation means 1101 and loaded as needed. In this case, the operation can quickly shift to subsequent processing.

20 Each processing executed in the embedding position determination means 103 will be described next in detail.

[3-2-1 Mask Generation Means]

The mask generation means 1101 will be described
25 first.

In embedding the additional information Inf using

the patchwork method, if the information is embedded by largely manipulating pixel values to increase the resistance against attacks (for example, when the value c in the pattern array is set to be large), degradation
5 in quality of the image represented by the original image data is relatively inconspicuous at a so-called edge portion where the pixel value abruptly changes. However, at a flat portion where the change in pixel value is small, the portion manipulated in its pixel
10 value becomes conspicuous as noise.

Fig. 29 is a graph showing the radial frequency characteristic perceived by the human eye. The abscissa represents the radial frequency, and the ordinate represents the visual response value. As is
15 apparent from Fig. 29, when pixel values are manipulated to embed information, degradation in image quality is conspicuous in the low-frequency region where the sensitivity of the human eye is high.

For this reason, in the second embodiment, the
20 pattern corresponding to each bit is laid out in consideration of the characteristics of a blue noise mask or cone mask normally used to binarize a multilevel image.

The characteristics of a blue noise mask and cone
25 mask will be briefly described.

First, the characteristics of a blue noise mask

will be described.

As a characteristic of a blue noise mask, a blue noise pattern is always obtained independently of the threshold value used for binarization. The blue noise pattern exhibits a frequency characteristic in which the radial frequency has a bias in the high-frequency region.

Fig. 53 is a view showing part of a blue noise mask.

Fig. 30A is a graph schematically showing the radial frequency characteristic of a blue noise mask binarized using a threshold value "10".

The abscissa in the graph shown in Fig. 30A indicates the radial frequency that represents the distance from the origin (DC component) for Fourier transform of the blue noise mask. The ordinate indicates a power spectrum which is a value obtaining by calculating and averaging the square-sum of amplitude components at distances represented by the radial frequencies on the abscissa. The graph of Fig. 30A one-dimensionally shows the two-dimensional frequency characteristic of an image to help understanding.

As compared to Fig. 29, since the blue noise mask has a bias in the high-frequency component and is therefore hardly perceived by the human eye. As is

00615577.071300
masks for all grayscale levels are generated to
generate a dither matrix.

For example, in a 32×32 blue noise mask, the
number of points increases (decreases) by four in units
5 of grayscale levels.

However, black (white) bits determined on the
basis of the previous grayscale level g cannot be
inverted to obtain 256 grayscale levels. For this
reason, restriction conditions become serious for a low
10 or high grayscale level, and only a nonuniform random
pattern is obtained.

Fig. 28 is a graph showing the appearance
frequency distribution (histogram) of coefficients of a
blue noise mask. Referring to Fig. 28, all values
15 (coefficients) "0" to "255" are present in the same
number in the mask.

A technique using a blue noise mask to binarize a
multilevel image is well known. This technique is
disclosed in detail in, e.g., Tehophano Mitsa, Kevin J.
20 Parker, "Digital halftoning technique using a blue
noise mask, J. Opt. Soc. Am A/Vol. 9, No. 11/November
1992.

The characteristics of a cone mask will be
described next.

25 As one characteristic feature of a cone mask,
when coefficients included in this mask are binarized,

00615577.074300

periodical or pseudo-periodical peaks are generated on a radial frequency region representing the resultant binary information, as shown in Fig. 30B. However, the cone mask is designed to have no peaks in the

5 low-frequency region.

Fig. 54 is a view showing part of the coefficient array of a cone mask.

Since an appropriate distance is maintained between dots independently of the threshold value used to binarize the cone mask, no peaks are generated in the low-frequency region.

Fig. 30B is a graph schematically showing the radial frequency characteristic of a cone mask binarized using a threshold value "10". Like the radial frequency characteristic of the blue noise mask shown in Fig. 30A, the characteristic shown in Fig. 30B also has a small number of low-frequency components.

In the cone mask, since peaks are generated from frequencies higher than the low frequency of a blue noise mask independently of whether the threshold value is large or small, the number of dense embedding positions is smaller than that in the blue noise mask. For this reason, embedded noise generated when the additional information Inf is embedded is more unnoticeable than blue noise.

The use frequency of coefficients of the cone

mask also exhibits the appearance frequency distribution (histogram) shown in Fig. 28, as in the blue noise mask.

When a pattern corresponding to each bit information of the additional information Inf is embedded in image data in correspondence with each coefficient of the mask, patterns equal in number to the bit information can be arranged in the image data. As a consequence, the embedded additional information Inf can be balanced.

In the second embodiment, a cone mask is used as an embedding reference mask because of the above advantages.

[3-2-2 Mask Reference Means]

The mask (cone mask) generated by the mask generation means 1101 is input to the mask reference means 1102.

The mask reference means 1102 makes the embedding positions of the N-bit information to be embedded in the image correspond to the mask numbers (pixel values), thereby determining the embedding positions.

A method of determining the embedding position by the mask reference means 1102 will be described.

In this embodiment, the above-described cone mask is used. For the descriptive convenience, a 4×4 mask 1501 shown in Fig. 31 is used.

The mask shown in Fig. 31 has 4×4 coefficients. That is coefficients "0" to "15" are laid out one by one. The embedding position of the additional information Inf is referred to using the 4×4 mask.

5 For the mask used for this description, the additional information Inf having 16 bits at maximum can be embedded. However, a description will be made below assuming that the additional information Inf having 8 bits is to be embedded.

10 The structure of the additional information Inf will be described first with reference to Fig. 52. As shown in Fig. 52, the additional information Inf is formed from start bits Inf_1 and use information Inf_2 .

The start bits Inf_1 are used by an offset
15 matching means included on the digital watermark extraction apparatus side to recognize the shift of the actual embedding position of the additional information Inf from an ideal position and accordingly correct the extraction start position of the digital watermark
20 (additional information Inf). This will be described later in detail.

The use information Inf_2 is used as the actual additional information, i.e., information to be actually used as additional information of the image
25 data I. For example, to track the cause for illicit use of the image data wI , the ID of the apparatus shown

In this embodiment, the ratio of the number of coefficients equal to or smaller than the threshold value used for embedding to the total number of coefficients in the mask will be called an embedding
5 filling rate.

To accurately embed the 8-bit additional information Inf an integer multiple number of times, the threshold value for determining a coefficient that is to be used as an embedding reference position in the
10 mask 1501 shown in Fig. 31 must be 8 or 16. As this threshold value, an optimum value is determined in consideration of the influence to the resilience and image quality.

When the threshold value of the mask is 8, the
15 embedding filling rate is 50%. That is, 50% of the original image data collated with the mask is subjected to processing using the pattern array shown in Fig. 25.

Table 1 shows an example of the correspondence between bit information and coefficients in a mask.

20 <Table 1>

Order of Bit Information To Be Embedded	S1	S2	S3	S4	S5	1	2	3
Coefficients in Mask	0	1	2	3	4	5	6	7

S1 to S5 are pieces of bit information (start bits) used by the offset matching unit for positioning, and 1 to 3 are three bits of use information.

According to the correspondence shown in Table 1, the pieces of bit information are embedded, using a pattern (Fig. 25), at the pixel positions in the input image data in correspondence with the position of coefficients (0 to 7) represented by 1601 in Fig. 32. The correspondence between the order of bit information to be embedded and the coefficient values in the mask is one of key information. Each bit information cannot be extracted without knowing the correspondence. In this embodiment, for the descriptive convenience, the start bits S1 to S5 and the three bits of use information are made to correspond to the coefficient values from 0 to the threshold value, as shown in Table 1.

The filling rate for actually embedding using a 32×32 cone mask will be briefly described next. The processing procedure is the same as in use of the mask 1501.

First, in consideration of degradation in image quality in embedding, a threshold value necessary for accurately embedding the additional information Inf an integer multiple number of times is determined.

To repeatedly embed the bit information of the additional information Inf in the same repetitive number of times, the number of coefficients equal to or smaller than the threshold value is divided the number

N of bits forming the additional information Inf, thereby determining the number of times for embedding each bit in one mask size.

For example, to embed the above-described 69-bit additional information Inf having five start bits and 64-bit use information in original image data corresponding to coefficient values 0 to 255, the threshold value is set to, e.g., 137.

In this case, the number of effective coefficient values in the mask is 138. Since the number of bits necessary for expressing one additional information Inf is 69, each bit information can be embedded twice ($= 138/69$) in one mask size.

In determining the embedding positions using the cone mask, pieces of bit information are embedded at all points with coefficient values equal to or smaller than a certain threshold value so as to exploit the characteristic of the cone mask in which no peaks are generated in the low-frequency components of the radial frequency.

When the embedding positions are determined in the above-described way, consequently, the embedding filling rate is 50%, and the embedded information amount is 69 bits. In this case, a relationship as shown in Table 2 holds between the bit information of the additional information Inf and the coefficient

values in the cone mask.

<Table 2>

Order of Bit In- formation To Be Em- bedded	S1	S2	S3	S4	S5	1	2	..	64
Coeffi- cients in Mask	0,1	2,3	4,5	6,7	8,9	10,11	12,13	..	136,137

S1 to S5 are start bits or bit information used by the offset matching unit for positioning, and 1 to 64 are bits of use information.

The present invention is not limited to this correspondence. Another correspondence may be set between the bit information and the coefficient values as long as the pieces of bit information are sequentially embedded, using the pattern shown in Fig. 25, at all positions of the coefficients from 0 to the threshold value (or from the threshold value to 255).

In a 32 x 32 cone mask, four positions with the same coefficient are present in one mask.

When pieces of bit information are embedded in the original image data in correspondence with all coefficients on the basis of Table 2 using a large cone mask such as a 32 x 32 or 64 x 64 cone mask, the pieces of bit information of the additional information Inf are embedded an almost equal number of times. Additionally, pieces of identical bit information are

Top660

09615577.07.1300

spread and embedded in the original image data.

In the patchwork method, embedding positions are conventionally selected such that patterns (Fig. 25) corresponding to the bit information do not overlap each other. In this embodiment, however, the same effect as described above can be obtained by referring to the cone mask. In addition, degradation in image quality is small.

As a result, the mask reference means 1102 obtains the coordinates (x,y) of the embedding position corresponding to each bit information.

This information is represented by array $S[\text{bit}][\text{num}] = (x,y)$, in which bit represents the start bits S1 to S5 and three bits of use information in Table 1, and num is the order of coefficients that repeatedly appear in the cone mask. The coordinates (x,y) represent relative coordinates in the mask.

The above operation is performed by the mask reference means 1102.

[3-2-3 Mask Pattern Array Corresponding Means]

The embedding position of each bit information in the cone mask, which is obtained by the mask reference means 1102, is input to the mask pattern array corresponding means 1103.

The embedding position determined by the mask reference means 1102 is the pattern position

(corresponding to 8×8 pixels) of the pattern of each bit information. In the patchwork method, addition regions (+c), subtraction regions (-c), and regions (0) except these regions shown in Fig. 25 must be assigned.

- 5 To do this, the mask pattern array corresponding means 1103 bitmaps the pattern array with an 8×8 size corresponding to Fig. 25 to all positions in the cone mask referred to by the mask reference means 1102.

- 10 More specifically, for coordinates represented by array $S[\text{bit}][\text{num}] = (x, y)$ obtained by the mask reference means 1102, the x-coordinate is multiplied by the horizontal size of the pattern array, and the y-coordinate is multiplied by the vertical size of the pattern array. As a consequence, the coordinates 1701
- 15 in the mask shown in Fig. 33A become the start coordinates 1702 for which one pixel in the mask shown in Fig. 33B is extended to one pattern array.

- When the pattern array shown in Fig. 25 is applied starting from the start coordinates, the bit
- 20 information can be embedded without overlapping a region 1703 having the pattern array size.

The coordinates (x, y) change to coordinates (x', y') , though bit and num of the array $S[\text{bit}][\text{num}]$ do not change.

- 25 Hence, (x', y') is defined as the start position at which the additional information Inf corresponding

to bit of the array S[bit][num] is embedded in accordance with the pattern array, so a plurality of bit information can be embedded.

The large mask obtained by bitmapping (enlarging) each coefficient of the cone mask to an 8×8 pattern array by the mask pattern array corresponding means 1103 is called an enlarged mask.

The size of the enlarged mask is $(32 \times 8) \times (32 \times 8)$. This size is the minimum necessary image unit (called a macro block) used to embed at least one additional information Inf.

The operation performed by the mask pattern array corresponding means 1103 has been described above.

A small mask generally has a lower degree of freedom in dot layout for mask generation than that of a large mask, so it is difficult to generate a mask such as a cone mask having desired characteristics. For example, when the additional information Inf is embedded by repeatedly assigning a small mask to the entire image data, the radial frequency of the small mask appears in the entire image data.

On the other hand, since the complete additional information Inf is extracted from one mask, the extraction resilience (possibility of extracting the additional information Inf from the partial image data wI') becomes low when a large mask size is set. For

this reason, the mask size must be determined in consideration of balance between the extraction resilience and the degradation in image quality.

Processing performed by the mask pattern array
5 corresponding means 1103 shown in Fig. 17 has been described above.

[3-3 Additional Information Embedding Processing]

The additional information embedding means 104
shown in Fig. 117 actually embeds the additional
10 information Inf by referring to the embedding position determined in the above way for each bit information in the image data.

Fig. 26 is a flow chart showing processing of repeatedly embedding the additional information Inf.

15 In the processing shown in Fig. 26, a plurality of assignable macro blocks are assigned to the entire image. In addition, the first bit information is repeatedly embedded in all of these macro blocks. Subsequently, the second bit information, third bit
20 information,... are repeatedly embedded. If unembedded bit information remains, processing by means 1001 to 1003 is executed for all unprocessed macro blocks.

However, the present invention is not limited to this sequence, and the relationship between the two
25 loop processes may be reversed. More specifically, when an unprocessed macro block remains, all bit

information unembedded in this macro block may be embedded.

More specifically, when bit information of the additional information Inf to be embedded is "1", the pattern array shown in Fig. 25 is added. If the bit to be embedded is "0", the pattern array shown in Fig. 25 is subtracted. That is, pattern arrays with positive and negative signs inverted from those in Fig. 25 are added.

The addition/subtraction processing is realized by selectively controlling the switching means 1001 shown in Fig. 26 in accordance with bit information to be embedded. More specifically, when bit information to be embedded is "1", the switching means 1001 is connected to the addition means 1002. When the bit information is "0", the switching means 1001 is connected to the subtraction means 1003. Processing by the means 1001 to 1003 is executed while referring to the bit information and pattern array information.

Fig. 35 is a view showing the process of embedding one bit information. In the example shown in Fig. 35, the bit information to be embedded is "1", i.e., the pattern array is added.

In the example shown in Fig. 35, $I(x,y)$ is the original image, and $P(x,y)$ is the 8×8 pattern array. The coefficients of the 8×8 pattern array are

superposed on the original image data (blue component) having the same size as that of the pattern array, and values at the same position are added/subtracted. As a result, $I'(x,y)$ is calculated and output to the color component synthesis means 105 shown in Fig. 17 as the image data of blue component in which bit information is embedded.

The above-described addition/subtraction processing using the 8×8 pattern array is repeatedly performed for all the embedding positions determined on the basis of Table 2 (positions at which the pattern array for embedding bit information is assigned).

Figs. 34A and 34B are views showing internal loop processing in Fig. 26.

Referring to Figs. 34A and 34B, macro blocks 1802 are repeatedly assigned and embedded (1001 to 1003 in Fig. 26) in an entire image data 1801 (1803) starting from the upper left corner to the lower right corner in accordance with the raster sequence so as to repeatedly embed each bit information.

The above operation is performed by the additional information embedding means 104, so the additional information Inf is embedded in the entire image.

With the above processing, the additional information Inf is embedded in the image data. If each

pixel of the image data having the additional information Inf embedded is represented by a sufficiently small number of dots, the size of the pattern array is also sufficiently small, and each
5 pattern array is perceived by the human eye as only a small dot. Hence, the radial frequency characteristic of the cone mask is also maintained unnoticeable to the human eye.

[3-4 Registration Processing]

10 The registration means 202 shown in Fig. 18, which is provided on the digital watermark extraction apparatus side, will be described next in detail.

The registration means 202 is located on the input side of the additional information extraction
15 means 203 for pre-processing of extraction processing of the additional information Inf. The image of blue component extracted by the color component extraction means 201 on the input side is input to the registration means 202.

20 The registration means 202 corrects the difference in scale between the image data wI output from the digital watermark embedding apparatus and the image data wI' input to the digital watermark extraction apparatus.

25 Fig. 23 is a view showing a detailed arrangement of the registration means 202. The registration means

202 comprises a block segmentation means 701, Fourier transform means 702, impulse extraction means 703, scaling rate calculation means 704, and scaling means 705.

5 The block segmentation means 701 performs the same block segmentation processing as that of the above-described registration signal embedding means 102 (block segmentation means 401). With this processing, it generally becomes difficult to extract the same
10 block as by the registration signal embedding means 102. This is because the image data wI with digital watermark information embedded is processed by the printing system to change its size and shift its position.

15 However, even when the blocks relatively inaccurately extracted, no problem is posed. This is because the digital watermark embedding apparatus has embedded a registration signal in the amplitude spectrum of the image data. As the nature of the
20 amplitude spectrum, it is not influenced by a positional shift in the spatial domain of the image data. Hence, no problem is posed even when the blocks segmented by the block segmentation means of each of the digital watermark embedding apparatus and digital
25 watermark extraction apparatus have slight positional shifts in the spatial domain.

The block segmentation means 701 outputs the image data segmented into blocks to the Fourier transform means 702. Like the above-described registration signal embedding means 102, the Fourier transform means 702 converts the image data in the spatial domain into image data in the frequency domain. The Fourier-transformed image data in the frequency domain is represented by an amplitude spectrum and phase spectrum. Only the amplitude spectrum is input to the impulse extraction means 703 while the phase spectrum is discarded.

The image data converted into the frequency domain is input to the impulse extraction means 703. The impulse extraction means 703 extracts only impulse signals from the image data converted into the frequency domain. More specifically, the impulse signals 502, 503, 504, and 505 shown in Fig. 21, which are already embedded in the image data, are extracted.

This processing can be performed using a known image processing technique. For example, this processing can be realized by threshold processing of the image data converted into the frequency domain. Fig. 24A shows this example. Fig. 24A is a graph showing a state wherein an amplitude spectrum 801 input to the impulse extraction means 703 is processed using a threshold value 802. For the descriptive convenience,

the converted image is one-dimensionally expressed in Fig. 24A. When the appropriate threshold value 802 is selected, the impulse signals can be extracted.

However, the original image data having almost the same
5 size as impulse signals present in the low-frequency region is also simultaneously extracted.

Fig. 24B is a graph showing the scheme of this embodiment that solves the above problem. The image data 801 converted into the frequency domain is
10 quadratically differentiated. This processing is equivalent to Laplacian filtering. Reference numeral 803 denotes data obtained by quadratically differentiating the image data 801 converted into the frequency domain. For this data 803, an appropriate
15 threshold value 804 is selected, and threshold processing is performed, thereby extracting an impulse signal.

Impulse signal extraction will be described with reference to Fig. 42 using a more detailed principle.
20 Fig. 42 also shows processing on the above-described registration signal embedding side.

The registration signal embedding means 102 converts image data 2601 in the spatial domain into image data 2602 in the frequency domain, and impulse
25 signals 2603 are added in the frequency domain.

The image data in the frequency domain, to which

the impulse signals 2603 are added, is subjected to inverse frequency conversion and returned to a signal 2601' in the spatial domain. The image data 2601' returned to the spatial domain should have the influence of addition of impulse signals. However, this influence is hardly perceived by the human eye, and the image data 2601 and 2601' almost look like the same data. This is because the impulse signals 2603 added in the frequency domain are distributed to the entire image at a small amplitude by inverse Fourier transform.

Adding the impulse signal 2603 as shown in Fig. 42 is equivalent to adding image data having a predetermined frequency component to the spatial domain. If the added impulse signal has a frequency higher than that perceivable by the human eye and an amplitude smaller than a limit perceivable by the human eye, the added impulse signal is invisible to the human eye. Hence, registration signal embedding processing is a kind of digital watermark processing.

In this embodiment, after the registration signal 2603 is embedded in the image data 2601, and the additional information Inf to be embedded is actually embedded, the signal 2601' in the spatial domain is reconstructed.

The registration signals embedded as shown in

Fig. 42 are Fourier-transformed again for extraction.
The registration signals 2603 temporarily spread to the
entire image in the spatial domain are converted into
the frequency domain and appear as impulse signals
5 again.

When an image with digital watermark information
embedded is attacked by, e.g., irreversible compression
such as JPEG compression, the amplitude of the impulse
signal becomes small at a high possibility. When the
10 image data receives a geometrical attack such as
scaling, the position of impulse signal moves. In any
case, the impulse signal can be extracted by
appropriate impulse extraction processing as described
above, and a change from the original image data can be
15 estimated. When this change is corrected, reliable
extraction of the additional information Inf embedded
in this embodiment is enabled.

With this above processing, the above-described
impulse signal is output from the impulse extraction
20 means 703 shown in Fig. 23 and input to the scaling
rate calculation means 704. The scaling rate
calculation means 704 calculates the type of scaling
using the coordinates of the received impulse signal.

In this embodiment, assume that the digital
25 watermark extraction apparatus side knows the frequency
component in which the impulse signal has been embedded

in advance. In this case, the scaling rate can be calculated on the basis of the ratio of the frequency component in which the signal has been embedded in advance to the frequency from which the impulse is
5 detected. For example, letting a be the frequency in which the impulse signal is embedded in advance, and b be the frequency of the detected impulse signal, it is found that scaling at a ratio a/b has been performed. This is a well-known nature of Fourier transform. With
10 the above processing, the scaling rate is output from the scaling rate calculation means 704.

However, the present invention is not limited to this, and information of the registration signal embedded position (frequency) may be received from the
15 digital watermark embedding apparatus side as needed. For example, the present invention also incorporates a case wherein the scaling rate is calculated upon receiving the position information as an encrypted signal. In this arrangement, only those who know the
20 registration signal can accurately extract the additional information Inf. In this case, the registration signal can be used as a key for extracting the additional information Inf.

The scaling rate output from the scaling rate
25 calculation means 704 is input to the scaling means 705. The image data wI_i is also input to the scaling means

the determined region using Table 2 and pattern array shown in Fig. 25.

The additional information Inf is extracted by convoluting the pattern array in the determined region.

5 [3-5-2 Reliability Distance Calculation Means]

The reliability distance d is a calculation value which is required to extract embedded information.

Fig. 22 is a flow chart showing a method of obtaining the reliability distance d corresponding to each bit information.

Processing executed by a convolution calculation means 601 shown in Fig. 22 will be described first with reference to Figs. 37 and 38.

Figs. 37 and 38 are views showing example in which one-bit information of the additional information Inf is to be extracted.

In the example shown in Fig. 37, 1-bit information of the additional information Inf is extracted from image data (blue component) $I''(x,y)$ in which the 1-bit information is embedded. In the example shown in Fig. 38, 1-bit information is extracted by way of trial from the image data $I(x,y)$ having no 1-bit information embedded.

Referring to Fig. 37, $I''(x,y)$ is image data with 1-bit information embedded, and $P(x,y)$ is the 8×8 pattern array used for convolution processing (pattern

array used to extract the additional information $I''(x,y)$. Each element $(0, \pm c)$ of the 8×8 pattern array is integrated with a pixel value arranged at the same position of the input image data $I''(x,y)$, and also, the
 5 sum of integrated values is calculated. That is, $P(x,y)$ is convoluted in $I''(x,y)$. In this case, $I''(x,y)$ is an expression including an image obtained when the image data $I'(x,y)$ has received an attack. If no attack is made, $I''(x,y) = I'(x,y)$. When 1-bit
 10 information is embedded in the image data $I''(x,y)$, a non-zero value is obtained at a very high possibility as a result of convolution calculation, as shown in Fig. 37. Especially when $I''(x,y) = I'(x,y)$, the convolution result is $32c^2$.

15 In this embodiment, the pattern array used for embedding and that used for extraction are the same. However, the present invention is not limited to this. Generally, when the pattern array used for embedding is $P(x,y)$, and the pattern array used for extraction is
 20 $P'(x,y)$, the relationship therebetween can be rewritten to

$$P'(x,y) = aP(x,y)$$

where a is an arbitrary real number. In this embodiment, a case wherein $a = 1$ will be described for
 25 the descriptive convenience.

In the example shown in Fig. 38, the same

00615577.071300

calculation as described above is performed for image data $I(x,y)$ without 1-bit information embedded. A zero value is obtained from the original image (corresponding to the image data I) as expectation of a
5 result of convolution calculation.

The method of extracting 1-bit information has been described above with reference to Figs. 37 and 38. However, in the above-described case, the result of convolution is 0 in the image data I in which the
10 additional information Inf is embedded. This is a very ideal case. However, in a region corresponding to the 8×8 pattern array of the actual image data I , the result of convolution calculation rarely becomes 0.

More specifically, when convolution calculation
15 is performed using the pattern array shown in Fig. 25 (the cone mask is also referred to as layout information) for a region corresponding to the 8×8 pattern array of the original image (image data I), a non-zero value may be undesirably calculated.
20 Conversely, when convolution calculation is performed for a region corresponding to the 8×8 pattern array of the image (image data wI) having the additional information Inf embedded, not the value $32c^2$ but "0" may be obtained.

25 However, each bit information of the additional information Inf is normally embedded in the original

image data a plurality of number of times. That is,
 the additional information Inf is embedded in the image
 a plurality of number of times. Hence, in n macro
 blocks having bit information embedded, convolution
 5 calculation is performed in units of 8×8 pattern
 arrays, and on the basis of the n results of
 convolution calculation for each bit information, it is
 statistically determined whether "each bit information
 is embedded" or "each bit information is 1 or 0". The
 10 statistic determination method will be described later.

The convolution calculation means 601 obtains the
 sum of the plurality of convolution calculation results
 for each bit information of the additional information
 Inf. For example, if the additional information Inf
 15 has eight bits, eight sums are obtained. The sum
 corresponding to each bit information is input to an
 average calculation means 602. Each sum is divided by
 the total number n of macro blocks and averaged. This
 average value is the reliability distance d. That is,
 20 the reliability distance d has a value generated by
 deciding by majority whether the value is similar to
 " $32c^2$ " or "0" in Fig. 37.

In the above-described patchwork method, however,
 the reliability distance d is defined as $d = 1/N \sum (a_i$
 25 $- b_i)$. Strictly, the reliability distance d is the
 average value of convolution results using $P'(x, y) =$

1/c $P(x,y)$. However, even when convolution calculation is performed using $P'(x,y) = aP(x,y)$, only a multiple of real number of the reliability distance d is obtained as the average value of the convolution calculation results, so the same effect as described above can be sufficiently obtained. Hence, in the present invention, the average value of convolution calculation results using $P'(x,y) = aP(x,y)$ can be sufficiently used as the reliability distance d .

10 The obtained reliability distance d is stored in a storage medium 603.

 The convolution calculation means 601 repeatedly generated the reliability distances d for the bit information of the additional information Inf and sequentially stores them in the storage medium 603.

 This calculated value will be described in more detail. The reliability distance d calculated for the original image data I using the pattern array shown in Fig. 25 (the cone mask is also referred to as layout information) is ideally 0. In the actual image data I , however, although this value is very close to 0, it is often non-zero. Fig. 39 is a graph showing the frequency distribution of the reliability distance d generated for each bit information.

25 Referring to Fig. 39, the abscissa represents the value of the reliability distance d generated for each

bit information, and the ordinate represents the number of bit information (the appearance frequency of the reliability distance d) for which convolution is performed to generate the reliability distance d . As is apparent from Fig. 39, this distribution is similar to the normal distribution. In addition, in the original image data I , although the reliability distance d is not always 0, the average value thereof is 0 (or a value very close to 0).

On the other hand, when the above convolution is performed for not the original image data I but the image data (blue component) $I'(x,y)$ in which the bit information "1" has been embedded, as shown in Fig. 35, the reliability distances d have a frequency distribution as shown in Fig. 40. That is, as shown in Fig. 40, the distribution shifts to the right while maintaining its shape in Fig. 39. As described above, in the image data in which a certain bit of the additional information I_{inf} is embedded, the reliability distances d are not always 0, though the average value thereof is c (or a value very close to c).

An example wherein the bit information "1" is embedded is shown in Fig. 40. When bit information "0" is embedded, the frequency distribution shown in Fig. 39 shifts to the left.

As described above, when the additional

information Inf is to be embedded using the patchwork method, the number of bits to be embedded (the number of times of pattern array use) is made as large as possible such that the statistic distribution as shown in Fig. 39 or 40 accurately appears. More specifically, whether bit information of the additional information Inf is embedded or whether the embedded bit information is "1" or "0" can be detected at a high accuracy.

[3-5-3 Offset Matching Processing]

10 The arrangement of an offset matching means 2002
will be described next.

The offset matching means 2002 receives the image data wI'_2 after appropriate scaling. After that, the start bits are detected using the reliability distance calculation shown in Fig. 22. The offset matching means 2002 generates only five reliability distances d corresponding to the five start bits Inf_i . The start bits Inf_i are part of the additional information Inf embedded in advance by the additional information embedding means 104 and comprise 5 bits in this embodiment, as shown in Fig. 52.

The start bits Inf_i comprise the first five bits as a concept. However, in an image having the additional information Inf embedded, the start bits are present not adjacently or densely but sparsely. This is because the pieces of bit information are

00615577.071300

point. The reason for this will be described with reference to Fig. 43.

Fig. 43 shows a state wherein to extract the additional information Inf embedded by the patchwork method used in the second embodiment, the same pattern array (2702 and 2705) as that used to embed the additional information Inf is used (the cone mask is also referred to as layout information) to search for the original macro block position (2701, 2703, and 2704) while performing convolution. Searching continuously progresses from the left to the right.

Referring to Fig. 43, one macro block (minimum unit with which the additional information Inf can be extracted) as part of the image data wI' , will be exemplified for the descriptive convenience. One cell in Fig. 43 represents the size of pattern array used to embed 1-bit information.

When the macro block 2701 and the pattern array 2702 have the relationship shown on the left side of Fig. 43, i.e., the pattern array 2702 is located on the upper left side of the actual macro block 2701, the pattern arrays for the original image and for additional information Inf extraction overlap only in the hatched regions.

At the center of Fig. 43, the position during searching and the actual macro block position

completely match. In this state, the pattern array to be convoluted and the macro block overlap at maximum.

On the right side of Fig. 43, the position during searching is located on the lower right side of the macro block position at which the additional information Inf_i is actually embedded. In this state, the pattern array to be convoluted and the macro block overlap only in the hatched regions.

If the pattern array to be convoluted and the macro block sufficiently overlap in all cases shown in Fig. 43, the correct start bits Inf_i can be detected. However, since the overlap area changes between the three cases, the reliability distance d also changes.

The overlap area can be replaced with the reliability distance d . More specifically, when the pattern array to be convoluted and the macro block completely match, the reliability distance d of each bit information becomes very close to the above-described $\pm 32c^2$.

In this embodiment, as shown in Fig. 44, if it is determined in step 2803 that the detected bits are not the correct start bits Inf_i , processing moves to the next search point in accordance with the raster sequence. If it is determined that the bits are the correct start bits Inf_i , it is determined in step 2804 whether the reliability distance d is larger than the

maximum value MAX. If NO in step 2804, the maximum value MAX is updated to the current reliability distance d , and the current search point is stored as the embedding start point. It is determined in step 5 2806 whether all search points have been searched. If NO in step 2806, processing moves to the next search point in accordance with the raster sequence. If YES in step 2806, the embedding start point stored at that time is output, and the processing is ended.

10 By the series of processing operations, the offset matching means 2002 of this embodiment detects the start bits Inf_i , determines, as the embedding start point of the additional information Inf , information of coordinates with the largest reliability distance d in 15 the coordinates at which the correct start bits Inf_i are obtained, and outputs the information to the output side as embedding start coordinates.

[3-5-4 Use Information Extraction Means]

A use information extraction means 2003 receives 20 the embedding start coordinates and image data with the additional information Inf embedded from the offset matching means 2002 on the input side, calculates the reliability distance d for only bit information of the use information Inf_i using the same operation as 25 described with reference to Fig. 22, and output reliability distances d_1 for the bit information to a

00615577.071300

statistical authorization means 2006 on the output side.

Obtaining the reliability distance $d1$ corresponding to each bit information of the use information Inf_1 almost corresponds to extracting each bit of the embedded use information Inf_1 . This will be described later.

At this time, only each reliability distance d is calculated on the basis of the embedding start coordinates determined by the above searching, and the five bits of the start bits Inf_1 are not extracted. [3-6 Statistical Authorization Processing]

The statistical authorization means 2006 determines the reliability of the reliability distance $d1$ obtained by the use information extraction means 2003 shown in Fig. 36. This determination is done by generating a reliability distance $d2$ using a second pattern array different from the first pattern array used to extract the additional information Inf (use information Inf_1) and generating the reliability index D by referring to the appearance frequency distribution of the reliability distance $d2$.

The reliability distance $d1$ is a reliability distance obtained by using the first pattern array (the cone mask is also referred to as layout information) in order to extract the use information Inf_1 by the use information extraction means 2003. The reliability

distance d_2 is a reliability distance obtained using the second pattern array (to be described later) different from the first pattern array. The first pattern array is normally the pattern array shown in Fig. 25, which is used to embed the additional information Inf (start bits Inf_1 and use information Inf_2).

The second pattern array and reliability index D will be described later in detail.

[3-6-1 Extraction Processing Using Second Pattern Array]

<<Central-Limit Theorem>>

$\{a_1, a_2, \dots, a_N\}$ and $\{b_1, b_2, \dots, b_N\}$ are sets of pixel values each consisting of n elements and correspond to pixel values of elements of the subset A and subset B as shown in Fig. 46.

When each of $\{a_1, a_2, \dots, a_N\}$ and $\{b_1, b_2, \dots, b_N\}$ has a sufficient number N of elements, the pixel values a_i and b_i have no correlation, and the expected value for the reliability distance d ($\sum (a_i - b_i)/N$) is 0. By the central-limit theorem, the reliability distances d exhibit an independent normal distribution.

The central-limit theorem will be briefly described.

In this theorem, when an arbitrary sample having a magnitude n_c is extracted from population (the

population need not always have a normal distribution) with an average value m_c and standard deviation σ_c , the distribution of average value S_c approaches a normal distribution $N(m_c, (\sigma_c / \sqrt{n_c})^2)$ as n_c becomes large.

5 Generally, the standard deviation σ_c of the population is often unknown. However, the number n_c of samples is sufficiently large, and the number N_c of population is sufficiently larger than the number n_c of samples, the standard deviation s_c of the sample may be
10 used in place of σ_c without posing any practical problem.

Referring back to this embodiment, the appearance frequency distribution of the reliability distances d_l obtained by the use information extraction means 2003
15 largely changes depending on whether the use information Inf_2 is accurately extracted.

In case of a detection error of the start bits Inf_1 (in case of an offset matching error), no bit information is actually embedded at the position where
20 the use information Inf_2 should be embedded. Hence, the appearance frequency distribution of the reliability distances d_l is given as a normal distribution 2501 shown in Fig. 41.

On the other hand, when the start bits are
25 accurately detected, the reliability distances d_l corresponding to bit information "0" of the use

the center to 95%) of the normal distribution 2501 generated using the reliability distances d2, the target image has a statistical bias. It can be suggested that the use information Inf_1 is embedded, so
5 the accuracy of the use information Inf_1 can be statistically determined. A method of this determination will be described later.

Next, a method of, using image data with the additional information Inf (use information Inf_1 ,
10 embedded), generating a distribution (normal distribution 2501 as shown in Fig. 41) similar to the appearance frequency distribution of the reliability distances d1 before the additional information Inf is embedded will be described.

15 In this embodiment, the reliability distances d2 which form a distribution similar to the normal distribution 2501 are obtained using an extraction means 2005 using the second pattern array.

The extraction means 2005 using the second
20 pattern array is a means for obtaining the reliability distance d2 using the second pattern array "perpendicular" to the first pattern array used by the use information extraction means 2003. The operation itself is almost the same as that of the use
25 information extraction means 2003, including convolution processing.

For a comparison, the pattern array shown in Fig. 25 used by the use information extraction means 2003 and the mask (cone mask) used to refer to the layout position of the pattern array will be called a "first pattern array" and "first position reference mask", respectively, and a pattern array "perpendicular" to the first pattern array and a mask used to refer to the layout position of the pattern array will be called a "second pattern array" and "second position reference mask", respectively.

The extraction means 2005 using the second pattern array receives the embedding start coordinates from the offset matching means 2002 and also calculates the reliability distance d2 using the above-described reliability distance calculation in Fig. 22.

The pattern array used for the reliability distance calculation in Fig. 22 is not a pattern array 901 shown in Fig. 25, which is used for embedding, but a pattern array 3601 or 3602 "perpendicular" to the pattern array 901.

This is because the influence of manipulation of the pattern array 901 shown in Fig. 25, which is used to embed the additional information Inf, is not reflected to the pattern arrays 3601 and 3602 in Figs. 49A and 49B.

As shown in Fig. 50, the result obtained by

convoluting the pattern array 901 shown in Fig. 25 and pattern array 3601 "perpendicular" to the pattern array is 0. This also applies to the pattern array 3602.

That is, the convolution result for the first and
5 second pattern arrays is 0. Hence, even when the density of the original image is changed using the first pattern array, this does not influence the reliability distance d obtained by convolution processing using the second pattern array.

10 The appearance frequency distribution of the reliability distances d_2 , which is obtained by performing convolution processing using the second pattern array on the image with the additional information Inf embedded is almost the same as the
15 normal distribution 2501 shown in Fig. 41. Hence, the appearance frequency distribution is regarded as the normal distribution 2501.

The resultant normal distribution 2501 is the criterion necessary for statistical authorization
20 processing 3507 in Fig. 48.

As described above, the extraction means 2005 using the second pattern array generates the normal distribution of reliability distances d_2 using the "pattern array perpendicular to the first pattern",
25 such as the pattern array 3601 or 3602 in Fig. 49A or 49B, and the second position reference mask 3802 shown

00615577.071300

in Fig. 51B.

Conditions of the "pattern array perpendicular to the first pattern array" are as follows.

(1) The size is the same as that of the pattern array 901 in Fig. 25, as shown in Figs. 49A and 49B.

(2) As in the pattern arrays 3601 and 3602, the result of convolution processing for the pattern array 901 in Fig. 25, which is used to embed the additional information Inf, is 0.

The convolution processing shown in Fig. 50 is the same as that shown in Figs. 37 and 38.

In this embodiment, that state wherein the convolution result becomes 0 is equivalent to the fact that the inner product of vectors becomes 0 when they are perpendicular to each other, and is expressed as "the pattern arrays are perpendicular to each other". Hence, the pattern array 3601 or 3602 in Fig. 49A or 49B is the "pattern array "perpendicular" to the pattern array 901 in Fig. 25".

The pattern array "perpendicular" to the pattern array used to embed the additional information Inf is used to calculate the reliability distance d2 because an appearance frequency distribution having no statistical bias in the distribution of the reliability distances d2, i.e., having 0 at the center is generated.

Another necessary condition of the "pattern array

"perpendicular" to the first pattern" is as follows.

(3) The pattern array has non-zero elements equal in number to non-zero elements of the pattern array used in the use information extraction processing 2003, and the number of positive elements equals the number of negative elements.

This aims at extracting the reliability distance d1 and reliability distance d2 under the same calculation conditions.

10 In this embodiment, the "second position reference mask" has a pattern different from that of a mask 3801 used to embed the additional information Inf and uses a reference mask 3802 shown in Fig. 51B, which has a size different from that of the mask 3801.

15 When the first and second pattern arrays are different, the appearance frequency distribution of the reliability distances d2 is almost the same as the normal distribution 2501.

20 However, if the start bit detection position is not accurate, a statistical bias may be detected even when convolution is performed using the second pattern array. In this embodiment, this possibility is also taken into consideration, and the periodical elements are canceled by making the first and second reference mask sizes different from each other. Alternatively, 25 convolution in the same region is not performed by

changing the pattern array layout in the mask.

In this case, the "second position reference mask" need not always be a cone mask as far as the coefficients of the mask are distributed at random.

5 To set the "second embedding position reference mask" different from the "first embedding position reference mask", the "second embedding position reference mask" is generated by an embedding position determination means 2004 in Fig. 36.

10 Generally, when the above-described extraction resilience is taken into consideration, the first position reference mask (cone mask) does not have a large size relative to the entire image data in which the additional information Inf is to be embedded.

15 Hence, a relatively large mask is preferably used as the "second position reference mask". In this embodiment, the size of the second mask used to calculate the reliability distance d1 on the additional information Inf side is set to be larger than that of
20 the first mask referred to in embedding the additional information Inf.

However, the present invention is not limited to this and can provide the effect to some extent even when the mask sizes are equal. Hence, the "second
25 position reference mask" may be generated by the embedding position determination means 2001 in Fig. 36.

00ET20" 2255T960

The minimum necessary condition for the masks is that the numbers of times of repeating each bit of the additional information Inf applied to the masks are equal in an image region with the same size.

5 If no sufficient result is obtained by extraction processing using the second pattern array, another second pattern array or second position reference mask having the above conditions is used to calculate the reliability distance d2 again. In this case, the
10 normal distribution 2501 shown in Fig. 41 may be generated as the ideal appearance frequency distribution.

The detailed operation of the extraction means
2005 using the second pattern array will be described
15 next.

In this embodiment, the first position reference mask is a 32 x 32 cone mask, and the second position reference mask is a 64 x 64 cone mask. The relative layouts of the coefficients between the two masks are
20 completely different.

In the extraction means 2005 using the second pattern array, the extraction position is determined in accordance with Table 3.

<Table 3>

Order of Bit Information	1	2	3	4	..	64
Coefficient Values in	0, 1	2, 3	4, 5	6, 7	..	136, 137

T1020

Second Position Reference Mask



In the second position reference mask, 16 coefficients with the same value are present in the mask. On the other hand, in the first position reference mask having a size of 32 x 32, one coefficient is repeated four times in the 32 x 32 size when the mask is referred to in the above-described Table 2. That is, in image data having the same size, the number of coefficients with the same value in the first position reference mask is equal to that in the second position reference mask.

In this embodiment, the second pattern array is assigned to a positional relationship according to the rule in Table 3, and convolution processing is sequentially executed to calculate 69 reliability distances d2 corresponding to the bit information. [3-6-2 Reliability Index D]

The reliability distances d2 generated by the extraction means 2005 using the second pattern array appear in almost the same distribution as the normal distribution 2501. In the normal distribution, it is known that 95% samples (reliability distances d2) generally appear within the range of the following inequality (25.1).

$$m - 1.96\sigma < d2 < m +$$

1.96 σ ... (25.1)

where σ is the standard deviation for the reliability distance d_2 , and m is the average.

The above range is called a "95% confidence
5 interval".

The values $m - 1.96\sigma$ and $m + 1.96\sigma$ are calculated using the reliability distance d_2 obtained by the second extraction means 2005 using the second pattern array.

10 The appearance frequency distribution of the reliability distances d_1 input from the use information extraction means 2003 to the statistical authorization means 2006 is the normal distribution 2502 shown in Fig. 41 when the bit information is "1" and the normal
15 distribution 2503 when the bit information is "0". For this reason, the reliability distance d_1 corresponding to the use information Inf_i is present outside the 95% confidence interval (hatched portion in Fig. 41) obtained by the extraction means 2005 using the second
20 pattern array at a very high probability.

At the time of processing by the offset matching means 2002, if the use information Inf_i is not present in the image to be processed, the appearance frequency distribution of the reliability distances d_1 is also
25 given as the normal distribution 2501.

In this case, all of the 64 reliability distances

d1 corresponding to the use information Inf_i , are not included in the confidence interval of inequality (25.1) at a probability as low as $(1 - 0.95)^{64}$.

Hence, when the normal distribution 2501 is
5 obtained on the basis of the reliability distance d2, it can be almost reliably determined whether the additional information Inf (use information Inf_i) is embedded by determining whether the appearance frequency distribution obtained on the basis of the
10 reliability distance d1 is included within the range that accounts for a greater part of the normal distribution 2501.

The statistical authorization means 2006 determines, using the above-described nature, the
15 reliability that the additional information Inf (use information Inf_i) is embedded.

In this embodiment, the reliability that the use information Inf is embedded is handled as the reliability index D.

20 The reliability index D is defined by the ratio of the number of reliability distances d1 outside the range of inequality (25.1) to all reliability distances d1 generated by the use information extraction means 2003.

25 If the reliability index D is larger than a threshold value α , the statistical authorization means

2006 determines that the overall appearance frequency distribution of the reliability distances d_1 is artificially biased to the position 2503 or 2503 in Fig. 41, i.e., the use information Inf , is properly embedded.

Hence, the statistical authorization means 2006 considers that the reliability distance d itself, which is used for determination, is reliable information and permits to further transfer the reliability distance d_1 to a comparison means 2007 on the output side.

As shown in the reliability display step 3510 in Fig. 48, the reliability index D of the use information Inf , or a message based on the index D may be displayed on a monitor or the like.

For example, when the reliability index D is not larger than the threshold value α , a message "the use information Inf , is not accurately extracted" is displayed, and the flow returns from the statistical authorization step 3507 in Fig. 48 to the step 3502 of inputting the image again.

[3-7 Comparison Processing]

The comparison means 2007 shown in Fig. 36 receives the value of reliability distance d_1 output through the use information extraction means 2003 and statistical authorization means 2006. Since the input reliability distance d_1 is reliably information, it

09615577-071300

need be only simply determined whether bit information corresponding to each reliability distance d_1 is "1" or "0".

More specifically, when the reliability distance
5 d_1 of given bit information of the use information Inf_2 has a positive value, it is determined that this bit information is "1". If the reliability distance d_1 has a negative value, the bit information is determined to be "0".

10 The use information Inf_2 obtained by the above determination is output as final data to be used as user reference information or control signal:

The series of processing operations from additional information embedding to extraction have
15 been described above.
(Modifications)

In the above embodiment, error-correction-coded additional information Inf (use information Inf_2) may be used. In this case, the reliability of the
20 extracted use information Inf_2 further improves.

The present invention may be applied as part of a system constituted by a plurality of devices (e.g., a host computer, an interface device, a reader, a printer, and the like) or to part of an apparatus comprising a
25 single device (e.g., a copying machine, a facsimile apparatus, or the like).

09616577.071300

The present invention is not limited to the apparatus and method for realizing the above embodiments. The present invention also incorporates a case wherein software program codes for realizing the
5 above embodiments are supplied to the computer (CPU or an MPU) in the system or apparatus, and the computer in the system or apparatus causes various devices to operate in accordance with the program codes, thereby realizing the above embodiments.

10 In this case, the software program codes realizes the functions of the above-described embodiments by themselves, and the present invention incorporates the program codes themselves and a means for supplying the program codes to the computer and, more particularly, a
15 storage medium storing the program codes.

As a storage medium for storing the program codes, a floppy disk, a hard disk, an optical disk, a magnetooptical disk, a CD-ROM, a magnetic tape, a nonvolatile memory card, a ROM, or the like can be used.

20 Not only in a case wherein functions of the above-described embodiments are realized when the computer controls various devices in accordance with only the supplied program codes but also in a case wherein the above embodiments are realized by the
25 program codes in cooperation with an OS (Operating System) running on the computer or another application

software, the program codes are incorporated in the present invention.

09615577.071300

The present invention also incorporates a case wherein the above embodiments are realized when the supplied program codes are stored in the memory of the function expansion board of the computer or function expansion unit connected to the computer, and the CPU of the function expansion board or function expansion unit performs part or all of actual processing on the basis of the instructions of the program codes.

In the above embodiments, digital watermark information is embedded using a cone mask. However, the present invention is not limited to this. Especially, the present invention also incorporates embedding digital watermark information using a blue noise mask.

In addition, an arrangement including at least one of the above-described various characteristic features is incorporated in the present invention.

As has been described above, according to the present invention, digital watermark information is embedded using the visual characteristics of a cone mask used for binarization. In embedding digital watermark information by partially adding/subtracting an image, the digital watermark information can be embedded while making degradation in image quality as

unnoticeable as possible to the human eye.

As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood
5 that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

09615577.071300